最適化理論

http://www.cs.miyazaki-u.ac.jp/~date/lectures/optimization/

伊達 章

宮崎大学 工学部 情報システム工学科

2016 年 5 月 27 日

講義のスケジュール（案）

1. 講義の概要
2. 数学的準備：曲線と曲面
3. 数学的準備：1 次形式と 2 次形式，2 次形式の標準形
4. 関数の極値
5. 関数の極値：ラグランジュの未定乗数法
6. 関数の最適化：勾配法・ニュートン法
7. 関数の最適化：共役勾配法
8. 統計的最適化：正規分布，最尤推定
9. 線形計画法 （その 1）
10. 線形計画法 （その 2）
11. 線形計画法 （その 3）
12. 非線形計画法
13. 動的計画法 （その 1）
14. 動的計画法 （その 2）
15. 定期試験，解説

目標

- 最適化手法の原理と計算法を学ぶ
  → 数理最適化ソルバー（ソフトウェア）を使いこなせるようになる
- 最適化理論に必要な数学
  - 曲線と曲面の方程式，法線ベクトル
  - 2 次形式の最大・最小化
- 関数の極値
  - ヘッセ行列やラグランジュの未定乗数法の意味
  - 最適化（勾配法など）
- 線形計画法の基本的な手法
- 動的計画法

テイラー展開： 苦手意識をなくそう

- 関数を多項式で展開

$$f(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \cdots$$

テイラー展開： 苦手意識をなくそう

- 関数を多項式で展開

$$f(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \cdots$$

何が面白い？

テイラー展開： 苦手意識をなくそう

- 関数を多項式で展開

$$f(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \cdots$$

何が面白い？

- $f(0) = c_0$
- $f'(x) = c_1 + 2c_2 x + 3c_3 x^2 + \cdots \to f'(0) = c_1$
- $f''(x) = 2c_2 + 6c_3 x + 12c_4 x^2 + \cdots \to f''(0) = 2c_2$
- $f'''(x) = 6c_3 + 24c_4 x + \cdots \to f'''(0) = 6c_3$

ある１点での詳細な情報 $\Rightarrow f$ の全体が復元できる！

テイラー展開： 苦手意識をなくそう

- 関数を多項式で展開

$$f(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \cdots$$

何が面白い？

- $f(0) = c_0$
- $f'(x) = c_1 + 2c_2 x + 3c_3 x^2 + \cdots \to f'(0) = c_1$
- $f''(x) = 2c_2 + 6c_3 x + 12c_4 x^2 + \cdots \to f''(0) = 2c_2$
- $f'''(x) = 6c_3 + 24c_4 x + \cdots \to f'''(0) = 6c_3$

ある１点での詳細な情報 $\Rightarrow f$ の全体が復元できる！

→ 本当か？

テイラー展開： 表現に２つある

A. $x = a$ を中心に，$f(x)$ を展開

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdot$$

$$f(x) = f(0) + f'(0)x + \frac{1}{2}f''(0)x^2 + \cdots$$

B. 上の式に $x = a + \Delta x$ を代入

$$f(a+\Delta x) = f(a)+f'(a)\Delta x+\frac{1}{2}f''(a)(\Delta x)^2+\cdots$$

- どう使い分ける？

テイラー展開： 表現に２つある

A. $x = a$ を中心に， $f(x)$ を展開

$$
\begin{aligned}
f(x) &= f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdot \\
f(x) &= f(0) + f'(0)x + \frac{1}{2}f''(0)x^2 + \cdots
\end{aligned}
$$

B. 上の式に $x = a + \Delta x$ を代入

$$
f(a+\Delta x) = f(a)+f'(a)\Delta x+\frac{1}{2}f''(a)(\Delta x)^2+\cdots
$$

- どう使い分ける？ → 結局は一緒． 少し先を知りたい
  A: p.57, 86, 89
  B: p. 2, 4, 11, 85, 88, 90

テイラー展開：2変数関数の場合

A. (a,b) を中心に，$f(x,y)$ を展開

$$f(x,y) = f(a,b) + \frac{\partial f(a,b)}{\partial x}(x-a) + \frac{\partial f(a,b)}{\partial y}(y-b) + \cdots$$

B. 上の式に $x = a + \Delta x, y = b + \Delta y$, を代入

$$f(a+\Delta x, b+\Delta y) = f(a,b) + \frac{\partial f(a,b)}{\partial x}\Delta x + \frac{\partial f(a,b)}{\partial y}\Delta y + \cdots$$

- 2変数の場合は1次近似で十分の場合が多い
  $(\Delta x)^2 \approx 0$, $\Delta x \Delta y \approx 0$

関数の最適化

3.1 勾配法

3.1.1 1 変数の場合

3.1.2 多変数の場合

3.2 ニュートン法： 収束が速い（2 次収束）

$f''(x)$ が計算できるなら，勾配法よりも効率的

2 次近似した放物線の極値を与える x を求める.

$\Rightarrow$ テイラー展開で 2 次近似して，$f'_{\mathrm{II}}(x) = 0$

3.2.1 1 変数の場合 $x_{t+1} \leftarrow x_t - \frac{f'(x_t)}{f''(x_t)}$

3.2.2 多変数の場合 $\boldsymbol{x}_{t+1} \leftarrow \boldsymbol{x}_t - H^{-1}(\boldsymbol{x}_t)\nabla f(\boldsymbol{x}_t)$

3.3 共役勾配法

3.3.1 2 変数の場合 $H(\boldsymbol{x}_t)\ (\boldsymbol{x}_{t+1} - \boldsymbol{x}_t) = \nabla f(\boldsymbol{x}_t)$

3.3.2 拡張と応用

関数の極値（復習）

- 第１章では，２次形式のみを扱った：

$$f(x,y) = \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}$$

- 第２章では，より一般的な場合を考える：

$$\begin{aligned}
f(x,y) &= ax + by + c \\
f(x_1,\cdots,x_n) &= \frac{1}{2}\sum_{i,j=1}^{n} h_{ij}x_ix_j + \sum_{i=1}^{n} a_ix_i + c \\
f(\boldsymbol{x}) &= \frac{1}{2}\boldsymbol{x}^{\mathrm{T}} H \boldsymbol{x} + \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} + c \\
\nabla f(\boldsymbol{x}) &= H\boldsymbol{x} + \boldsymbol{a} \\
H &= \nabla\nabla^{\mathrm{T}} f(\boldsymbol{x})
\end{aligned}$$

関数の極値

- 一般的な場合： $f = f(x_1, x_2, \cdots, x_n)$
  最大・最小値を求めることは不可能

- 局所的（$\leftrightarrow$ 大局的）な最大・最小を考える
  - 2 次関数の最小値
    ヘッセ行列の固有値がすべて正
  - 極値の判定（シルベスタの定理 p.52）
    対称行列が正値対称行列である必要十分条件
  - 関数 $f(\boldsymbol{x})$ を 2 次近似（テイラー展開）

$$f(\boldsymbol{x}) \approx \bar{f} + \sum_{i=1}^{n} \frac{\partial \bar{f}}{\partial x_i}(x_i - \bar{x}_i) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 \bar{f}}{\partial x_i \partial x_j}(x_i - \bar{x}_i)(x_j - \bar{x}_j)$$

  - 極値では

$$\frac{\partial \bar{f}}{\partial x_i} = 0, \ \ i = 1, \cdots, n$$

問題：ニュートン法 ***

次の関数を考える.

$$f(\boldsymbol{x}) = f(x, y) = 100(y - x^2)^2 + (1 - x)^2$$

・この関数の $\boldsymbol{x}_0 = (0, 0)^{\mathrm{T}}$ における 2 次近似

$$f_{\mathrm{II}}(\boldsymbol{x}) = \bar{f} + \nabla \bar{f} \cdot \boldsymbol{x} + \boldsymbol{x}^{\mathrm{T}} \bar{H} \boldsymbol{x}$$

を求めよ.

・その 2 次近似 $f_{\mathrm{II}}(\boldsymbol{x})$ が極値をとる点を求めよ.

共役勾配法の考え方

- 関数 $f(\boldsymbol{x})$ の $\boldsymbol{x}_0$ における 2 次近似

$$f_{\mathrm{II}}(\boldsymbol{x}) = \bar{f} + \nabla \bar{f} \cdot (\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0)^{\mathrm{T}} \bar{H} (\boldsymbol{x} - \boldsymbol{x}_0)$$

- $f_{\mathrm{II}}(\boldsymbol{x})$ が極値をとる点 $\boldsymbol{x}$ の必要条件

$$\nabla f_{\mathrm{II}}(\boldsymbol{x}) = \nabla \bar{f} + \bar{H}(\boldsymbol{x} - \boldsymbol{x}_0) = 0$$

- 現在値：$\boldsymbol{x}_0$. $\boldsymbol{m} = (\boldsymbol{x} - \boldsymbol{x}_0)$ 方向に移動 $\Rightarrow$ $\boldsymbol{x}$ に到達！
- $\boldsymbol{x}_0$ での接線 $\boldsymbol{t}$ と $\nabla \bar{f}$ は直交. $\boldsymbol{m} = \nabla \bar{f} + \alpha \boldsymbol{t}$ と表現.

$$(\boldsymbol{t}, \nabla \bar{f}) = (\boldsymbol{t}, -H\boldsymbol{m}) = -(\boldsymbol{t}, H(\nabla \bar{f} + \alpha \boldsymbol{t})) = 0$$

$$= \boldsymbol{t} H \nabla \bar{f} + \alpha \boldsymbol{t}^{\mathrm{T}} H \boldsymbol{t} = 0 \rightarrow \alpha = \frac{\boldsymbol{t} H \nabla \bar{f}}{\boldsymbol{t}^{\mathrm{T}} H \boldsymbol{t}}$$

終